

Original Paper

doi [10.15826/recon.2020.6.3.015](https://doi.org/10.15826/recon.2020.6.3.015)

JEL G18, H50, J11



COVID-19 mortality rate in Russian regions: forecasts and reality

M.L. Lifshits , N.P. Neklyudova*Institute of Economics of the Ural Branch of the Russian Academy of Science, Ekaterinburg, Russia; lifmarina@yandex.ru***ABSTRACT**

Relevance. COVID-19 is an extremely dangerous disease that not only spreads quickly, but is also characterized by a high mortality rate. Therefore, prediction of the number of deaths from the new coronavirus is an urgent task. **Research objective.** The aim of the study is to provide a more accurate estimate of the real number of coronavirus-related deaths in Russian regions. **Data and methods.** The main research method is econometric modeling. Comparison of various data was also applied. The authors' calculations were based on Rosstat data, the data of the World Bank and specialized sites with coronavirus statistics in Russia and in the world. **Results.** We identified the factors affecting the COVID-19 mortality rates in various countries were identified, assessed how much the official Russian statistics underestimated mortality in Russian regions, and provided predictive estimates of mortality as a result of the pandemic. We also determined the number of additional coronavirus-induced deaths. **Conclusions.** The official data on COVID-19 mortality in Russia underestimate the actual numbers more than twofold. The number of direct and indirect victims of the pandemic in Russia at the end of July was approximately 43 thousand people.

KEYWORDS

COVID-19, regions of Russia, mortality rate, lethality, econometric modeling

FOR CITATION

Lifshits, M.L., & Neklyudova, N.P. (2020) COVID-19 mortality rate in Russian regions: forecasts and reality. *R-economy*, 6(3), 171–182. doi: 10.15826/recon.2020.6.3.015

Смертность от COVID-19 в России и ее регионах: прогнозы и реальность

М.Л. Лифшиц , Н.П. Неклюдова*Институт экономики Уральского отделения Российской академии наук, Екатеринбург, Россия; lifmarina@yandex.ru***АННОТАЦИЯ**

Актуальность. COVID-19 – крайне опасное заболевание, которое характеризуется не только очень быстрым распространением, но и очень высокой летальностью. Поэтому прогноз числа смертей от COVID-19 является весьма актуальной задачей. **Целью работы** является более точная по сравнению с официальной оценка реального числа смертей от COVID-19 в России. **Данные и методы.** В работе использованы два основных метода исследования, эконометрическое моделирование и сравнение данных из различных источников (таблиц Росстата и специализированных сайтов со статистикой по коронавирусу в России и в мире). **Результаты.** Выявлены факторы, влияющие на уровень летальности COVID-19 в различных странах, выполнена оценка занижения смертности от COVID-19 в регионах РФ и построены прогнозные оценки смертности в результате эпидемии. Определено число дополнительных смертей из-за коронавируса. **Выводы.** Официальные данные о смертности в России занижены более чем в два раза. Число прямых и косвенных жертв эпидемии в РФ на конец июля составило примерно 43 тысячи человек.

КЛЮЧЕВЫЕ СЛОВА

COVID-19, регионы России, уровень смертности, летальность, эконометрическое моделирование

ДЛЯ ЦИТИРОВАНИЯ

Lifshits, M.L., & Neklyudova, N.P. (2020) COVID-19 mortality rate in Russian regions: forecasts and reality. *R-economy*, 6(3), 171–182. doi: 10.15826/recon.2020.6.3.015

Introduction

The COVID-19 outbreak began in Russia a month later than in Italy and Spain, but it didn't help avoid the negative scenario. By the end of April, Russia ranked ninth in terms of the number of cases (overtaking China); starting from

17–20 May, Russia was second, surpassed later by Brazil and then India. Starting from July 5th, Russia was fourth (Fig. 1). In May, Russia ranked second behind the US in terms of active cases (all cases minus diseased and recovered), and from mid-April to mid-May it was second after the US

in terms of new weekly cases. All figures in the curves show a rolling average during 7 days.

According to official data, COVID-19 death rate in Russia is not as bad as the spreading of the disease. Russia ranked 12th in terms of the total number of deaths from COVID-19 in early autumn (solid black curve in Fig. 2). The top 10 countries with the highest cumulative number of coronavirus-related deaths include the United States, Brazil, India and Mexico, with the United Kingdom, Italy, France and Spain occupying

5–8 places although the peak of deaths there was observed in early April and followed by Peru and Iran. As our study will show, the official Russian figures show a vast underestimation. Our estimates, which we believe to be more accurate, are shown in Fig. 2 with a black dotted line. In this case, Russia ranked 5th in the total number of COVID-19 related deaths, ahead of the UK. The countries in the legends in Figures 1 and 2 are shown in descending order of the total number of new coronavirus infections and COVID-19

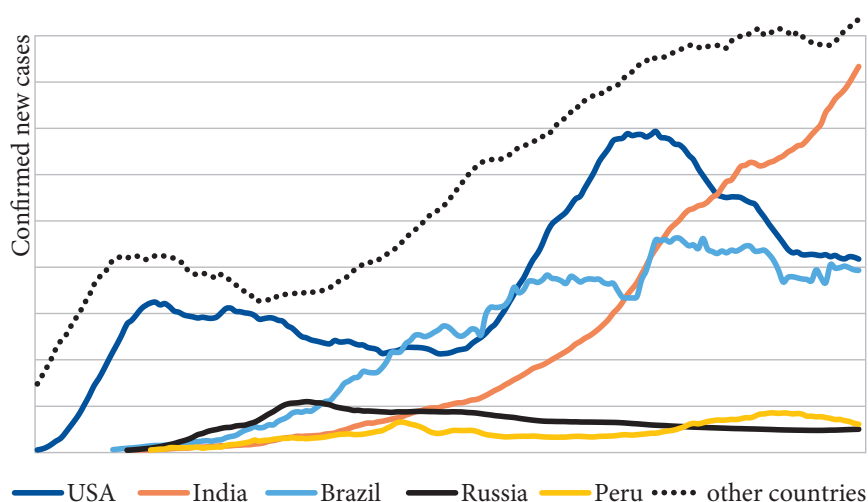


Figure 1. Daily new coronavirus cases in the countries with the highest number of confirmed cases (7-day moving average)

Source: the authors' calculations based on the data from WORLDOMETERS statistical data <https://www.worldometers.info/coronavirus/> (Accessed data: September 6, 2020)

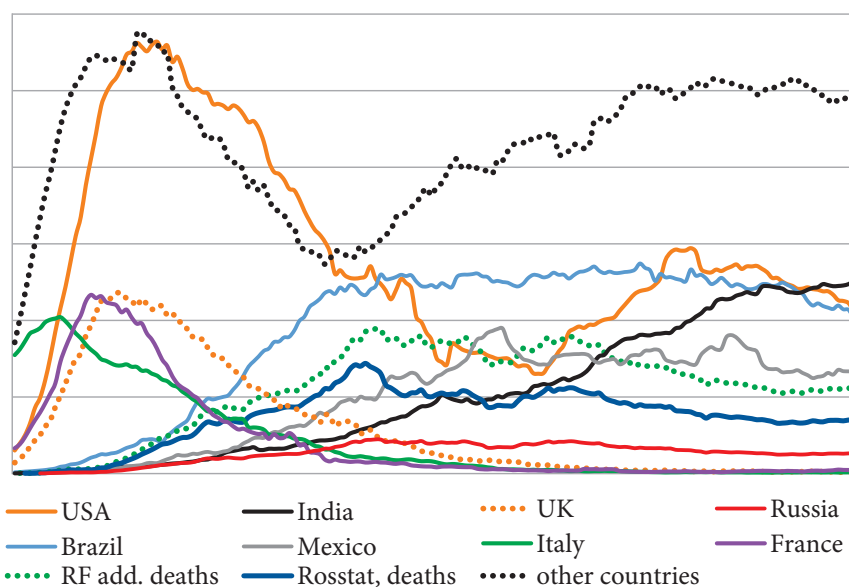


Figure 2. Daily COVID-19 deaths in the countries with the highest cumulative number of coronavirus-related deaths

Source: the authors' calculations are based on the data from WORLDOMETERS statistical data <https://www.worldometers.info/coronavirus/> and Federal State Statistics Service (Rosstat) <https://rosstat.gov.ru/folder/12781> (Accessed data: September 6, 2020)

deaths, respectively. It is also possible that in fact there are even more coronavirus-related deaths in Russia and that the curve “RF additional deaths” is the closest to the reality. In this case, Russia is in the third place in the world in terms of the total number of deaths. A detailed method for calculating the “Rosstat, deaths” and “RF additional deaths” curves is given at the end of the article.

In April, COVID-19 seemed to be an extremely dangerous disease, not only because it was spreading fast (the weekly number of the new cases in the world is still growing and in the last week of July reached its maximum), but also because of a high mortality rate. On April 10th–12th, the share of deaths among the registered cases in the world exceeded 22%, then gradually declined and by the beginning of September dropped below 4%. Even in recent months, the mortality rate still remains ten times higher than from influenza. Therefore, predicting the number of deaths from the new coronavirus was an urgent task in April and May.

A number of modelling and forecasting tools quantifying the future COVID-19 burden are available in different countries (Aronov et al., 2020; Drapkina et al., 2020; Ferraro et al., 2020; Haghani et al., 2020; Harapan et al., 2020; Scafetta, 2020; Singh et al. 2020). Sebastiani et al. (2020), Chintalapudi et al. (2020), Onder et al. (2020) described the trends of Covid-19 spread in Italy. Anastassopoulou et al. (2020) and Gao et al. (2020) revealed the potential total numbers of COVID-19 deaths in China. Shojaee et al. (2020) estimated the rate of deaths in Italy, Iran and South Korea. Sanchez-Villegas & Daponte Codina (2020) studied the COVID-19 epidemic in Spain. Semenova et al. (2020) forecast the number of COVID deaths in Kazakhstan. Gerli et al. (2020) tried to forecast mortality trends in 27 countries of the European Union, in Switzerland and the UK. Al-Raei (2020) calculated the coefficient of mortality of the COVID-19 pandemic for China, the United States, Russia, and the Syrian Arab Republic. Vандoros (2020) studied the impact of COVID-19 on mortality in England and Wales. Zemtsov & Baburin (2020) described the situation in Russia.

The purpose of the study is to offer a more accurate estimation of the real number of COVID-19-related deaths in Russia. To this end, we identified factors affecting COVID-19 mortality rates in various countries, evaluated the underestimation of COVID-19 mortality in Russian regions and made predictive estima-

tions of mortality as a result of the epidemic. We used the available data (Russian and international) to calculate new, more accurate estimates, while each next assessment of the new data that appeared also served as a verification of the previous forecast.

Methodology and data

We calculated the first COVID-19 mortality projections in Russia by using the data from other countries as of May 7th (Lifshits, 2020)¹. They were based on the statistical data available on the *worldometer* website. We also used the World Bank data (for the age structure of the population in countries of the world). All data on the pandemic on this site come from official sources of the countries concerned. Official data on Russia are provided by the Federal Service for Supervision of Consumer Protection and Welfare (Rospotrebnadzor), but the source of the data is not quite clear. The most significant disadvantage of these statistics is that the COVID-19 mortality data from different countries are not comparable (Danilova, 2020; Middelburg & Rosendaal, 2020; Methodological Recommendations, 2020). In some, for example, in Italy, all dead infected people are classified as COVID-19 victims, in others criteria are different. Although research shows that the new coronavirus causes a non-respiratory disease, its mechanism is much more complex (Varga et al., 2020), therefore, the Italian approach is more correct. On April 16th 2020, the WHO published guidelines, stating that “a death due to COVID-19 is defined as a death resulting from a clinically compatible illness, in a probable or confirmed COVID-19 case, unless there is a clear alternative cause of death (e.g. trauma). There should be no period of complete recovery from COVID-19 between illness and death. A death due to COVID-19 may not be attributed to another disease (e.g. cancer) and should be counted independently of preexisting conditions that are suspected of triggering a severe course of COVID-19”². However, even after the publication of these recommendations, they were interpreted differently across the world.

¹ Lifshits M. Mortality Statistics in Russia in Terms of Coronavirus. <https://promdevelop.ru/science/statistika-smertnosti-v-rossii-v-usloviya-koronavirusa/> (Accessed data: September 5, 2020)

² International guidelines for certification and classification (coding) of COVID-19 as cause of death. https://www.who.int/classifications/icd/Guidelines_Cause_of_Death_COVID-19.pdf (Accessed data: September 6, 2020)

However, even with these important limitations, the econometric modeling can trace some patterns in COVID-19 mortality rate (*proportion of deaths from completed cases*) in different countries. This indicator in percentage was taken as an explainable variable.

The econometric regression equation has the following form:

$$Y = b_0 + b_1 \cdot x_1 + \dots + b_n \cdot x_n + \varepsilon, \quad (1)$$

where Y is the explained variable, x_p , regressions, b_p , coefficients of regressions, ε , the remainder of the equation.

The econometric models for COVID-19 mortality predictions in Russia were based on data from 107 countries and territories out of 208, which were selected according to the following criteria:

1) the total number of cases – at least 450 (by the end of May 7, there were 122 such countries);

2) available data on testing – the Worldometer website provides the data on testing (9 dropped out) and the number of recoveries (3 more removed);

3) available data on the population's age structure – the World Bank website contains this information (3 more are excluded from consideration).

The following variables were taken as regressions:

1) Tests / infection – the number of tests per diagnosed infected person;

2) Completed, % – the percentage of completed cases (recovery plus death) among all infected;

3) 65 + / 15 – the percentage of the population aged 65 and over from the population aged 15 and over;

4) Date_250 is a dummy variable that takes the value of 1 for South Korea and adds 1 for each day that this country remained behind others in passing the threshold of 250 registered cases (for example, for Italy this figure is 3 and for Russia 28);

5) Tests / Population – the number of tests per thousand of the country's population.

The Memory Lists of Russian Deceased Health Workers³ compiled by their colleagues appeared in early May. At that time, it was the most accurate data on COVID-19-related mortality in the regions, since each deceased health worker is known by name. A significant drawback of these data was that at that time only in some regions there were several medical workers who died from coronavirus, and isolated cases could be a casual

occurrence. The comparison of the List with official data made it possible to establish that at the beginning of May, official mortality data were underestimated. This confirmed the conclusions drawn from the analysis of the remainder of the econometric equation described above.

In mid-July, Rosstat published preliminary data on births and deaths in Russian regions in April. These data, along with the Memory List, were used to estimate the real COVID-19 mortality in April and by mid-July. We also identified the regions where mortality in April was most underestimated. Rosstat receives its demographic information from the Unified State Register of Civil Registry, and there is no reason to doubt its adequacy. However, at that time there was a problem of significant inadequacy of the data related to the epidemic. The updated data of Rosstat for April, together with the preliminary data for May, were published on July 10; then, data for June and updated data for May appeared on August 7th, which we used to make new estimates and forecasts. Rosstat data for July appeared on September 4. Starting with the data for April, an additional table 5.1 "Information on the Number of Registered COVID-19 Deaths" appeared on Rosstat's website. Figures in this table can be compared with 'additional' deaths in the same period of the previous year. Thus, in addition to the least squares econometric modeling, the study compared the data from various sources and various Rosstat tables.

Results

COVID-19 mortality forecast in the Russian Federation by May 7th based on the data from the countries of the world

The data on 270 thousand COVID-19 deaths in 107 countries were used to study the factors affecting the mortality rate of COVID-19 in various countries with 148 thousand deaths recorded before the publication of WHO recommendations, when each country followed its own rules.

In Table 1, where the coefficients of the two econometric models are presented (variable to be explained and regressions are described above), the following notations are used: Constant – free term of the equation; Adj. R² is the corrected coefficient of determination; N is the number of observations (countries). The coefficients are random variables, so for each coefficient its standard deviation (standard error) is stated in parentheses.

³ Anonymous. Memory List. <https://sites.google.com/view/covid-memory/home> (Accessed data: September 6, 2020)

The level of coefficients' significance is indicated by asterisks, this is the probability that we consider an insignificant regressor significant: * –0.1; ** –0.5; *** –0.01; **** –0.001; the absence of asterisks means that the probability is more than 10%.

The models show that the proportion of completed cases has the greatest influence on the detected mortality rate, because at the epidemic's beginning, the most severe cases are usually noticed first.

The next most important factor is the age structure of the population. It should be noted that here the factor is the proportion of people 65+, not of the entire population but only of adults, since children, being the infection carriers, along with everyone else, rarely have clinical manifestations of the disease. In Russia, as a rule, the increased mortality rate is seen not where the general population is older, but where the proportion of older people in the adult population is higher. For example, in southern republics the population is relatively young, because the share of children is high, but the share of older people in the adult population is also higher than the national average, since life expectancy is higher. In Russia, the value of the index "65+/15+" (17.9) is significantly lower than in Italy (26.3) and Spain (22.7), therefore, lethality is expected to be lower.

The regressor "Date_250" is included in the model with a minus sign, because although there are still no reliable drugs for COVID-19 or a treatment protocol, the global medical community has been gaining experience, positive and negative.

In the second model, two factors characterizing the testing level are significant: per one case and per thousand people, both are only 10%, although the importance of testing is undeniable. Model 1 is presented to show that in the absence of one of these regressors, the significance of the

other is higher: about 5%. However, it is still lower than the significance of some other factors.

There are several reasons for this. First of all, different countries use different tests of different quality, and at the beginning of the pandemic there were no high-quality tests at all. Over time, all countries strove to increase testing per case in order to establish the presence or absence of the virus in as many people as possible who may have been in contact with the infected. However, from April 27th to May 7th, this indicator decreased in 27 countries and territories out of 107. In 5 of them this happened because the number of tests for some reason was not updated on the site (Egypt, Algeria, Guatemala, Mali, Norman Islands). It can be assumed that in the remaining 22, the number of cases is growing faster than the number of completed tests, because either the quality of the initial testing was especially low, or time was lost and the situation got out of control. The list (in decreasing order of this index decrease in absolute value) is the following: UAE, Ghana, Russia, Kenya, Bahrain, Colombia, Belarus, Nigeria, Bolivia, Brazil, Chile, Honduras, Mayotte, Qatar, Afghanistan, Pakistan, Mexico, Peru, Armenia, Niger, Singapore, and Bangladesh. Thus, Russia is among the three worst countries. It is also possible that in some of these countries only a small share of the population is repeatedly tested, while testing of the rest of the population is not given due attention.

Let us now consider which observations in Model 2 have the largest and smallest residuals with respect to the standard deviation and try to understand the reasons.

Only Honduras (3.63) goes beyond the three sigma, followed by Belgium (2.93), Sweden (2.22), Hungary (2.16), Bolivia (2.11), Ecuador (1.89), Philippines (1.86), France (1.77), and Indonesia

Table 1

Regressors explaining the differences in COVID-19 mortality in different countries

Regressors	Model 1		Model 2	
	Coefficients	Std. Error	Coefficients	Std. Error
Constant	21.36****	(3.76)	22.19****	(3.76)
Tests/Infected	–0.05070**	(0.02398)	–0.04386*	(0.02412)
Completed cases, %	–0.1995****	(0.0347)	–0.1934****	(0.0345)
65+/15+	0.3233***	(0.1224)	0.3452***	(0.1220)
Data_250	–0.1403**	(0.0649)	–0.1581**	(0.0651)
Tests/Population			–0.05395*	(0.03209)
Adj. R2	0.351		0.362	
N	107		107	

Source: the authors' calculations based on statistical data of Worldometer website <https://www.worldometers.info/coronavirus/> and World Bank <https://www.worldbank.org/en/who-we-are/news/coronavirus-covid19> (Accessed date: September 6, 2020).

(1.65). The largest negative balances in absolute value are less in absolute value than positive ones: Singapore (–2.31), Qatar (–1.84), Japan (–1.75), Belarus (–1.65), Russia (–1.51), Saudi Arabia (–1.40), Serbia (–1.28), Chile (–1.28), Armenia (–1.19).

Curiously, 8 of these 18 countries were previously listed among the countries with possibly the biggest testing problems. Inadequate testing could lead to both an increase in mortality (Honduras, Bolivia) and an underestimation of the number of real deaths (Russia, Belarus, Chile, Qatar, Armenia, Singapore).

Large positive deviations from the values calculated according to Model 2 could also be caused by both real problems (for example, complete refusal of Sweden from quarantine, mortality outbreaks in nursing homes in France) and by peculiarities of the statistical accounting (for example, in Belgium, to COVID-19 deaths were also added deceased in nursing homes who were suspected of having the disease, even if there was no confirmation). As for the large negative deviations, one should remember that Japan more than once has been praised by the press for its good organization of the measures against the pandemic. Apparently, it was not in vain, because anyone in Japan can go to a local clinic and have a CT scan made without any problems.

But do Belarus and Russia have the same level of healthcare? Here it is appropriate to turn to the question of the proportion of medical workers among the infected and among the deceased infected. All over the world, health workers are at risk because they come into contact with infected people, while nowhere in the world at the beginning of the pandemic there was a sufficient amount of protective equipment. However, the situation with the proportion of the deceased infected in the world is fundamentally different, because health workers have the opportunity of early disease detection and timely treatment. In many countries, memory lists of deceased health workers are being created and this allows us to make a comparison. In Italy, the proportion of medical workers among the deceased infected is 0.6%, in Germany 0.2%, in the United States 0.3%, that is, the share of medical workers among the deceased is usually 5–15 times smaller. In Russia and Belarus, however, things are different. There were 14 names on the Belarusian Memory List as of May 5th, that is, the share of registered deaths from the pandemic was about 13%. On May 10th,

there were 147 names on the Russian List in 33 regions of the Russian Federation (Dyer, 2020), this is 7.7% of all COVID-19 victims in Russia, of which 47 were in Moscow (4.4% of the deceased), 27 in Dagestan (150%) (!!!), 21 in the Moscow region (11%), 14 in St. Petersburg (26%), 5 in the Krasnodar Krai (23%), the remaining 33 deceased health workers accounted for 5.9% of the remaining 563 deceased infected. This indicates both the underestimation of the COVID-19 deceased data in the Russian Federation and the unsatisfactory quality of the Russian healthcare system.

Thus, the relatively low proportion of COVID-19 deaths in Russia and Belarus is, obviously, primarily the result of a special approach to the cause of death determination. For example, while cancer, atherosclerosis or diabetes complicates the course of the disease resulting from the new coronavirus infection, then in most countries COVID-19 is indicated as the cause of death or one of the causes, and in Russia it is customary to indicate only one main cause, and in this case it is usually cancer, acute vascular disease or diabetes⁴. In essence, such definitions of the death cause are contrary to the WHO recommendations.

So, to estimate the real number of deaths from COVID-19 in Russia until May 8th and to predict the total number of victims of the pandemic in Russia, the models from Table 2: Model 2 in the first case and Model 1 in the second case are quite suitable.

Model 2 showed that the data on COVID-19 mortality in Russia are underestimating the death toll in the country by about 2.12 times. Therefore, the real number of deaths, according to Model 2, as May of 8th 2020, was approximately 3,440. The forecast estimate was based on the following considerations. Typically, the number of detected infections declines slower than increases. Therefore, if we assume that by May 10th, Russia had reached or almost reached the peak of the detected cases per day, then the total number of infected during the pandemic will vary from 750 thousand to a million people. Then the total number of victims of the pandemic can be 14–19 thousand people. However, the standard error of the equations is large due to large errors in the initial data and a small sample. Therefore, this forecast was very approximate.

⁴ New coronavirus infection (COVID-19). Rules of work of pathology departments. Temporary recommendations. Version 2. Moscow, April 27, 2020, pp. 23–28. http://patolog.ru/sites/default/files/metodichka_dz.pdf (Accessed date: September 6, 2020).

Estimation of the “additional” mortality rate during the pandemic based on Rosstat’s preliminary demographic data for April

Rosstat published data on fertility and mortality rates in Russian regions for April on June 13th with a note that the data may be incomplete due to the quarantine measures since not all Russians were able to process documents on the births and deaths of their family members.

In January and February 2020, due to the warm winter, the total number of deaths in Russia (excluding the Crimea) was 4.9% and 3.7% less than a year earlier. An excess of 0.6% was recorded in March, with a total decline of 2.8% in January–March. However, in a number of regions, this figure exceeded even the data for April, which were incomplete.

Only four regions saw an increase in the mortality rate in February, March, and April compared to the last year: in Moscow (by 190, 169 and 1934, respectively), Moscow region (258, 217 and 1018), Penza region (25, 87 and 52) and Chuvashia (71, 47, 34). In total, in March and April, the excess of the mortality rate was recorded in 19 regions: Leningrad region (by 369), Tomsk region (244), Kemerovo region (229), Lipetsk region (179), Nizhny Novgorod region (176), Sverdlovsk region (142), Yaroslavl region (90), Kostroma region (40), Tambov region (38), Samara region (17), the Republic of Tatarstan (164), Karelia (68), Altay (32), Buryatia (31), Udmurtia (29), Khabarovsk region (120), St. Petersburg (93), the Khanty-Mansiysk Autonomous District (85) and Yamalo-Nenets Autonomous District (16).

Obviously, not all of these additional deaths were associated with the coronavirus, especially in February and March. However, the connection between the April data and the pandemic is obvious: the correlation coefficient of mortality in April changes with the number of infected is 0.789, and with the number of Covid-19 deaths – 0.765. It should be noted that in all the cases, the correlation with the number of cases is higher than with the number of Covid-19 deaths. This is extra evidence that the official coronavirus mortality statistics are not very close to reality.

Let us find out what the real mortality from coronavirus could be in April and in which regions the underreporting of deaths in April in Rosstat data could be the most significant. To this end, first, we will construct two regression equations and find their residuals:

$$\Delta D_{apr} = -101.5 + 0.040 \cdot (Inf_{apt} - Rec_{mar}) + \varepsilon \quad (2)$$

$$\text{and } \Delta D_{apr} = -100.8 + 28.53 \cdot D_{med} + \varepsilon, \quad (3)$$

where ΔD_{apr} is the change in the mortality rate in April 2020 compared to April 2019; Inf_{apt} is the number of people infected with the new coronavirus by April 30th in regions; Rec_{mar} is the number of Covid-19 recoveries by March 31st; D_{med} is the number of infected health workers who died before May 10th, ε is the equations residual

The variables coefficients were found empirically from 83 observations, the determination coefficient is $R^2 = 0.624$ for the first equation and 0.368 for the second.

The number of infected deceased health workers adds to the idea of the epidemiological situation in the regions, since their share in the total number of infected people is higher in places where the severe cases proportion is higher among the infected, and the medical care organization is the least satisfactory. Undoubtedly, it would be better to include in the equation the number of infected health workers who died during April, but the Russian Memorial List does not specify the death dates, and we do not have the data for the period before May 10th.

The first equation states that, on average, 4 out of every 100 infected people died in April, while the second suggests that 28 other people died for every infected deceased health worker.

Unfortunately, the equations do not show how much the mortality rate in April is underestimated, since the sum of all values calculated from the equation is always equal to the sum of the initial values of the explained variable. However, if we assume that the majority of additional deaths in February–April in Moscow and the Moscow region are associated with the pandemic, then in these two regions the proportion of doctors among the deceased infected is only 2%. In other regions, this share is probably higher, since the state of the healthcare system is worse. Thus, the coefficient on the variable “Health worker deaths” in equation 2 appears to be close to reality. Therefore, the total number of additional deaths in April should be close to 4.5 thousand, and the total number of deaths from all causes was probably about 3% higher in April 2020 than in April 2019.

According to equation (2), mortality is most underestimated in Krasnodar Krai (residual is –1040), Dagestan (–563), Irkutsk (–387) and Rostov (–338) regions, Stavropol Krai (–261), Sverdlovsk (–228), Ulyanovsk (–190), Tula (–157)

regions, Altay Territory (–135), Saratov region (–134), Zabaykalsk region (–132), Pskov region (–131) and Chechen Republic (–118).

According to equation (3) – in Dagestan (–1278), Krasnodar Krai (–1143), Irkutsk (–383) and Rostov (–330) regions, Stavropol Krai (–296), Ulyanovsk (–203) and Sverdlovsk (–200) regions, St. Petersburg (–185), Tula Region (–148), Zabaykalsk region (–128), the Chechen Republic (–128) and Pskov region (–126).

To exclude any doubt, we can also compare the changes in birth and mortality rates in regions in January–March and in April in comparison with the similar periods of the previous year, because the problems with the registration of births and deaths were most likely in the same regions. Then it turns out that, perhaps, in most regions in April, not all births and/or deaths were recorded. In particular, this applies to such regions as the Republics of Dagestan, Ingushetia, Kalmykia, Chechen Republic, Pskov, Ulyanovsk, Irkutsk and Sakhalin regions, Krasnodar Krai and Zabaykalsk region. For more complete analysis of mortality in Russian regions in April see the study by Lifshits (2020).

Estimation of COVID-19 mortality based on demographic data from the RSSS for April–May, June–July

In preliminary data, the number of births in April 2020 in Russia (excluding the Crimea) amounted to 91.9% in 2019, and the number of deaths was 97.2%, and according to the updated data, 94.7% and 98.2% respectively. The mortality rate in April was adjusted in 8 regions: the republics of Dagestan (55.4), Ingushetia (38.1), Chechen (20.2), North Ossetia (12.1), Kabardino-Balkaria (8.4), Adygea (7.5), Karachay-Cherkessia (6.3) and Krasnodar Krai (5.7). At the same time, in the preliminary data for May, the number of births was only 90.1% of 2019, and the number of deaths is 112.4%. Perhaps a larger number of deaths in April than births was recorded in the statistics for May. The transfer of a part of births and deaths to the statistics of the next month, apparently, happened later. For example, a part of deaths in June in the Republics of Dagestan, Karachay-Cherkessia and North Ossetia were clearly transferred to July statistics, which can be seen from Table 2. Table 2 shows the ratio of the number of deaths in April, May, June and July to the same months of

Table 2
Ratio of birth and mortality rates in April, May, June, July and January–July in Russian regions to the same months of the previous year

Region	Number of deaths, 2020 in % in 2019				
	April	May	June	July	January–July
1	2	3	4	5	6
Ingushetia	112.7	169.4	129.5	192.0	125.3
Dagestan	104.3	159.4	109.4	178.1	121.8
Yamalo-Nenets Autonomous District	111.4	137.6	159.7	137.4	120.5
Chechen Republic	91.9	144.9	135.8	173.4	119.6
Moscow	119.5	157.2	141.6	107.1	116.1
Chuvashia	102.7	113.0	148.6	143.0	115.1
Tatarstan	98.7	104.0	138.7	154.2	114.8
Leningrad region	105.9	129.5	149.3	113.7	114.6
Moscow region	113.4	144.1	133.1	109.8	114.5
St. Petersburg	102.2	146.7	140.8	122.3	114.5
Penza region	103.3	126.4	137.1	128.1	114.4
Khanty-Mansiysk Autonomous District	97.1	107.4	129.0	134.8	110.9
Mordovia	87.9	141.3	121.3	126.2	109.5
Karachayevo-Cherkessia	96.4	107.9	96.2	174.1	108.9
Kamchatka	92.4	94.0	157.9	125.9	108.4
Samara region	96.5	105.0	124.0	139.2	108.2
Tula region	89.1	119.2	140.6	108.8	107.9
Bashkortostan	99.5	109.7	109.1	147.7	107.8
Nenets Autonomous District	87.9	87.8	122.7	125.0	107.8
North Ossetia	99.2	142.5	72.1	157.2	107.2

Source: the authors' calculations based on statistical data of the Federal State Statistics Service (Rosstat) <https://rosstat.gov.ru/folder/12781> (Accessed data: September 5, 2020)

the previous year in the 20 regions with the largest ratio of the number of deaths in January–July 2020 to 2019 (the 15 largest values for each month are indicated in bold, not all the largest values for individual months are included in the table).

Table 2 shows that in some regions the epidemic is subsiding (Moscow, Moscow Region, St. Petersburg), while in others it has not started to abate yet (Tatarstan, Bashkortostan, Samara Region, Khanty-Mansiysk Autonomous District).

Since later data with the accumulated result turn out to be more accurate than the data for previous months, it is better to take the sum with the accumulated result in econometric modeling: either all deaths that, according to Rosstat, are associated with coronavirus, or all additional deaths.

If we take the changes in the sum of the deaths number in April and May as the explained variable, the econometric equations will have the following forms:

$$\Delta D_{(4+5)} = -32.5 + 0.046 \cdot (Inf_5 - Rec_{mar}) + \varepsilon, \quad (4)$$

$$\Delta D_{(4+5)} = -95.1 + 77.0 \cdot D_{med(5)} + \varepsilon \quad (5)$$

and

$$\Delta D_{(4+5)} = -88.0 + 0.024 \cdot (Inf_5 - Rec_{mar}) + 43.3 \cdot D_{med(5)} + \varepsilon, \quad (6)$$

where $\Delta D_{(4+5)}$ is the change in the mortality rate in April and May 2020 compared to April–May 2019; Inf_5 is the number of people infected with the new coronavirus by May 31st in the regions; Rec_{mar} is the number Covid-19 recoveries by March 31st; $D_{med(5)}$ is the number of infected health workers who died before June 1st; ε is the equations residual.

The variables' coefficients were found empirically from 83 observations, the determination coefficient is $R^2 = 0,831$ for equation (4), 0,842 for equation (5), and 0,912 for equation (6). Thus, the relationship between excess mortality and the selected regressors is now much greater than it was according to the incomplete April data.

Using the information about recovered and infected people according to official data, as well as knowing the number of dead infected health workers (as of July 31, there were 620), it can be calculated by formula (6) that by the end of July, the total number of direct and indirect victims of the pandemic in Russia amounted to approximately 43 thousand people.

Fortunately, not all the regions have the doctors who died during the fight against the pandemic, so two regressors in equation (6) com-

plement each other. Now, it became possible to construct model (6) with both regressors, because the correlation of each of the regressors with the explained variable is greater than with each other. Perhaps this happened because now the information on the number of people infected in the regions is less reliable than on the number of “additional” deaths. This is also indicated by the comparison of the remainders of equations (4) and (5). Apparently, the record holder in terms of underestimating the number of cases in Russia is Dagestan, since it has the smallest residual in equation (5), –2788, and one of the largest residues in equation (4).

If we compare the dynamics of the case and death number in Russia and other countries, the fact that official Russian statistics started to underestimate the number cases in May becomes evident.

Comparing figures 1 and 2, it is easy to note that in the USA, Brazil and the world as a whole, the level of mortality is noticeably decreasing. This is particularly evident in France (Fig. 3), where, by the end of August, the storm in newly detected infections was even greater than it was at the end of March, but this did not lead to the increase in the number of deaths from COVID-19. Obviously, in recent months, mostly mild cases of the disease have been recorded there.

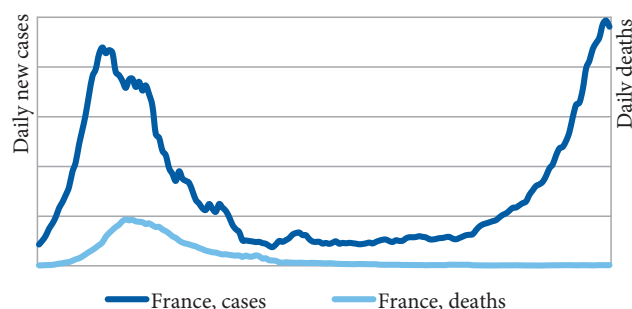


Figure 3. Daily new COVID-19 cases and COVID-19 deaths in France

Source: Lifshits M. Mortality statistics in Russia in terms of coronavirus. <https://promdevelop.ru/science/statistika-smertnosti-v-rossii-v-usloviya-koronavirusa/> (Accessed data: September 5, 2020)

However, the picture is fundamentally different in Russia (Fig. 4). The mortality from COVID-19 began to grow immediately after the President canceled the so-called “coronavirus holidays”. Obviously, since the President’s speech on May 10th, the official data have underestimated not only the number of COVID-19 related deaths, but also the number of new cases.

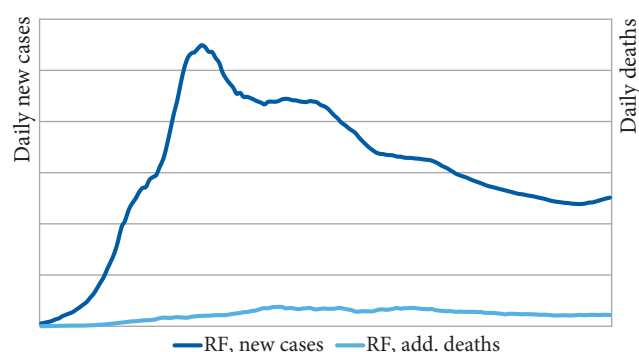


Figure 4. Daily New COVID-19 Cases and COVID-19 Deaths in Russia

Source: the authors' calculation based on the Federal State Statistics Service (Rosstat) <https://rosstat.gov.ru/folder/12781> and website CORONAVIRUS (COVID-19) <https://coronavirus-monitor.ru/coronavirus-v-rossii/> (Accessed data: September 6, 2020)

In Fig. 4, Rosstat data on additional deaths have been used as an estimate of the number of deaths from the pandemic since the beginning of the year, since by the end of July they significantly exceeded the number of coronavirus-related deaths: 58 thousand and 37.5 thousand, respectively. Since we do not know exactly how these additional deaths were divided into months in reality, to build the curve “Russia, additional deaths”, the official number of deaths was multiplied by the coefficient equal to the ratio of the total number of additional deaths for 7 months to the total number of deaths from COVID-19 by the end of July according to official data.

In Fig. 2, another curve was used with the estimation of the number of deaths in the Russian Federation: “Rosstat, deaths”. Different recalculation coefficients were used there for different periods: the official number of deaths before the end of May was multiplied by the ratio of the number of deaths from COVID-19 by the end of May according to Rosstat, to the official number of deaths before the end of May, the coefficient for June was the ratio of numbers for June, while the

coefficient for all subsequent data is the ratio of numbers for July.

After the appearance of Rosstat data for July, it became clear that the curve “Russia, additional deaths” more accurately reflects the reality.

Conclusion

The study used two main methods, econometric modeling and comparison of data from various sources and various Rosstat tables. Both methods showed identical results.

The econometric analysis of world data on May 7th concluded that official data in Russia underestimate the actual mortality rates more than twice. Our analysis of the Russian data for the same period which appeared later confirms this supposition. In general, each following forecast, as new data appeared, showed an increasingly distressing scene, and reality was ultimately even worse than forecasts. The reasons were, apparently, that the quality of testing in the Russian Federation was much worse than in developed countries, and after May 10th, the official data began to underestimate not only the number of COVID-19 deaths, but also the number of diseased.

Thus, our first assessment based on the analysis of world data showed approximately 3,440 deaths by May 8, and the estimation based on preliminary data for April – 4.5 thousand by the end of April. The forecast estimate based on preliminary data from Rosstat for May amounted to 43 thousand deaths by the end of July, and in reality there were 58 thousand additional deaths by this time. If the decline in the number of deaths in July (according to official data) continues, then the total number of additional deaths per year may approach 100 thousand. Now ranked third in the world in terms of the number of the pandemic victims, at the end of the year Russia will possibly share the 3rd or 4th place with India. It remains to be hoped that at least this forecast will not be overly optimistic.

References

- Al-Raei, M. (2020). The forecasting of COVID-19 with mortality using SIRD epidemic model for the United States, Russia, China, and the Syrian Arab Republic. *Aip Advances*, 10(6). doi: [10.1063/5.0014275](https://doi.org/10.1063/5.0014275)
- Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *Plos One*, 15(3). doi: [10.1371/journal.pone.0230405](https://doi.org/10.1371/journal.pone.0230405)
- Aronov, I.Z., Maksimova, O.V., & Galkina, N.M. (2020). COVID-19 Highest Incidence Forecast in Russia Based on Regression Model. *International Journal of Mathematical Engineering and Management Sciences*, 5(5), 812–819. doi: [10.33889/ijmems.2020.5.5.063](https://doi.org/10.33889/ijmems.2020.5.5.063)

Chintalapudi, N., Battineni, G., & Amenta, F. (2020). COVID-19 virus outbreak forecasting of registered and recovered cases after sixty day lockdown in Italy: A data driven model approach. *Journal of Microbiology Immunology and Infection*, 53(3), 396–403. doi: [10.1016/j.jmii.2020.04.004](https://doi.org/10.1016/j.jmii.2020.04.004)

Danilova, I.A. (2020). Morbidity and mortality due to COVID-19. The problem of data comparability. *Demographic review*, 7 (1), 6–26.

Drapkina, O.M., Samorodskaya, I.V., Sivtseva, M.G., Kakorina, E. P., Briko, N.I., Cherkasov, S.N., Zinserling V.A., & Malkov P.G. (2020). COVID-19: urgent questions for estimating morbidity, prevalence, case fatality rate and mortality rate. *Cardiovascular Therapy and Prevention*, 19(3):2585. (In Russ.) doi: [10.15829/1728-8800-2020-2585](https://doi.org/10.15829/1728-8800-2020-2585)

Dyer, O. (2020). Covid-19: Cases rise in Russia as health workers pay the price for PPE shortage. *Bmj-British Medical Journal*, 369. doi: [10.1136/bmj.m1975](https://doi.org/10.1136/bmj.m1975)

Ferraro, O.E., Puci, M.V., Montomoli, C., Rolesu, S., Cappai, S., & Loi, F. (2020). Official Data and Analytical Forecasts: Differences and Similarities Among Coronavirus Disease (COVID-19) Confirmed Cases and Deaths. *Frontiers in Medicine*, 7. doi: [10.3389/fmed.2020.00239](https://doi.org/10.3389/fmed.2020.00239)

Gao, Y., Zhang, Z., Yao, W., Ying, Q., Long, C., & Fu, X. (2020a). Forecasting the cumulative number of COVID-19 deaths in China: a Boltzmann function-based modeling study. *Infection Control and Hospital Epidemiology*, 41(7), 841–843. doi: [10.1017/ice.2020.101](https://doi.org/10.1017/ice.2020.101)

Gerli, A. G., Centanni, S., Miozzo, M. R., Virchow, J. C., Sotgiu, G., Canonica, G. W., & Soriano, J. B. (2020a). COVID-19 mortality rates in the European Union, Switzerland, and the UK: effect of timeliness, lockdown rigidity, and population density. *Minerva medica*. doi: [10.23736/s0026-4806.20.06702-6](https://doi.org/10.23736/s0026-4806.20.06702-6)

Haghani, M., Bliemer, M.C.J., Goerlandt, F., & Li, J. (2020a). The scientific literature on Coronaviruses, COVID-19 and its associated safety-related research dimensions: A scientometric analysis and scoping review. *Safety Science*, 129. doi: [10.1016/j.ssci.2020.104806](https://doi.org/10.1016/j.ssci.2020.104806)

Harapan, H., Itoh, N., Yufika, A., Winardi, W., Keam, S., Te, H., Megawati, D., Hayati, Z., Wagner, A.L., & Mudatsir, M. (2020). Coronavirus disease 2019 (COVID-19): A literature review. *Journal of Infection and Public Health*, 13(5), 667–673. doi: [10.1016/j.jiph.2020.03.019](https://doi.org/10.1016/j.jiph.2020.03.019)

Lifshits, M.L. (2020). *Regularities of the COVID-19 mortality level in the world and forecast of the number of death due to epidemic in Russia*. Paper presented at the XI Ural Demographic Forum, Yekaterinburg, Institute of Economics UB RAS, RUSSIA. (In press).

Methodological Recommendations on the Coding and Selection of the Underlying Medical Condition in the Statistics of Morbidity and the Primary Cause of Death in the Statistics of Mortality in Relation to COVID-19 (Approved by the Ministry of Health of the Russian Federation on 05.27.2020) (2020). *Russian Journal of Forensic Medicine*, 6(2), 53–62. (In Russ.) doi: [10.19048/2411-8729-2020-6-2-53-62](https://doi.org/10.19048/2411-8729-2020-6-2-53-62)

Middelburg, R.A., & Rosendaal, F.R. (2020). COVID-19: How to make between-country comparisons. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases*, 96, 477–481. doi: [10.1016/j.ijid.2020.05.066](https://doi.org/10.1016/j.ijid.2020.05.066)

Onder, G., Rezza, G., & Brusaferro, S. (2020). Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *Jama-Journal of the American Medical Association*, 323(18), 1775–1776. doi: [10.1001/jama.2020.4683](https://doi.org/10.1001/jama.2020.4683)

Sanchez-Villegas, P., & Daponte Codina, A. (2020). Predictive models of the COVID-19 epidemic in Spain with Gompertz curves. [Modelos predictivos de la epidemia de COVID-19 en Espana con curvas de Gompertz.]. *Gaceta sanitaria*. doi: [10.1016/j.gaceta.2020.05.005](https://doi.org/10.1016/j.gaceta.2020.05.005)

Scafetta, N. (2020). Distribution of the SARS-CoV-2 Pandemic and Its Monthly Forecast Based on Seasonal Climate Patterns. *International Journal of Environmental Research and Public Health*, 17(10). doi: [10.3390/ijerph17103493](https://doi.org/10.3390/ijerph17103493)

Sebastiani, G., Massa, M., & Riboli, E. (2020). Covid-19 epidemic in Italy: evolution, projections and impact of government measures. *European Journal of Epidemiology*, 35(4), 341–345. doi: [10.1007/s10654-020-00631-6](https://doi.org/10.1007/s10654-020-00631-6)

Semenova, Y., Glushkova, N., Pivina, L., Khismetova, Z., Zhunussov, Y., Sandybaev, M., & Ivanov, A. (2020). Epidemiological Characteristics and Forecast of COVID-19 Outbreak in the Republic of Kazakhstan. *Journal of Korean Medical Science*, 35(24). doi: [10.3346/jkms.2020.35.e227](https://doi.org/10.3346/jkms.2020.35.e227)

Singh, R.K., Rani, M., Bhagavathula, A.S., Sah, R., Rodriguez-Morales, A.J., Kalita, H., Nanda, C., Sharma, S., Datt Sharma, Y., Rabaan, A. A, Rahmani, J., & Kumar, P. (2020). Prediction of the COVID-19 Pandemic for the Top 15 Affected Countries: Advanced Autoregressive Integrated Moving Average (ARIMA) Model. *JMIR*

Shojaee, S., Pourhoseingholi, M.A., Ashtari, S., Vahedian-Azimi, A., Asadzadeh-Aghdaei, H., & Zali, M.R. (2020). Predicting the mortality due to Covid-19 by the next month for Italy, Iran and South Korea; a simulation study. *Gastroenterology and hepatology from bed to bench*, 13(2), 177–179.

Varga, Z., Flammer, A. J., Steiger, P., Haberecker, M., Andermatt, R., Zinkernagel, A.S., Mehra, M.R., Schuepbach, R.A., Ruschitzka, F., & Moch, H. (2020). Endothelial cell infection and endotheliitis in COVID-19. *Lancet*, 395(10234), 1417–1418. doi: [10.1016/s0140-6736\(20\)30937-5](https://doi.org/10.1016/s0140-6736(20)30937-5)

Zemtsov, S.P., & Baburin, V.L. (2020). Risks of morbidity and mortality during the COVID-19 pandemic in Russian regions. *Population and Economics*, 4(2), 158–181. doi: [10.3897/popecon.4.e54055](https://doi.org/10.3897/popecon.4.e54055)

Vandoros, S. (2020). Excess mortality during the Covid-19 pandemic: Early evidence from England and Wales. *Social Science & Medicine*, 258, 4. doi: [10.1016/j.socscimed.2020.113101](https://doi.org/10.1016/j.socscimed.2020.113101)

Information about the authors

Marina L. Lifshits – Postgraduate Student, Institute of Economics, Ural Branch of RAS (29, Moskovskaya Street, Yekaterinburg, 620014, Russian Federation) ; e-mail: lifmarina@yandex.ru

Natalia P. Neklyudova – Candidate of Sciences (Economics), Senior Researcher, Institute of Economics of the Ural Branch of the Russian Academy of Sciences (29, Moskovskaya Street, Yekaterinburg, 620014, Russian Federation) ; e-mail: nnp81@mail.ru

ARTICLE INFO: received July 15, 2020; accepted September 3, 2020

Информация об авторах

Лифшиц Марина Лазаревна – аспирант, Институт экономики Уральского отделения Российской академии наук (620014, г. Екатеринбург, ул. Московская, 29); e-mail: lifmarina@yandex.ru

Неклюдова Наталья Павловна – кандидат экономических наук, старший исследователь, Институт экономики Уральского отделения Российской академии наук (620014, г. Екатеринбург, ул. Московская, 29); e-mail: nnp81@mail.ru

ИНФОРМАЦИЯ О СТАТЬЕ: дата поступления 15 июля 2020 г.; дата принятия к печати 3 сентября 2020 г.